Report on the 2013 Assessments of Aurora and Rougheye Rockfishes

by R.I.C. Chris Francis
123 Overtoun Terrace
Wellington 6021
New Zealand

August 2013

Prepared for

The Center for Independent Experts
Northern Taiga Ventures, Inc.

Executive Summary

A STAR (stock assessment review) panel met 8-12 July 2013 at the Northwest Fisheries Science Center of NOAA/NMFS in Seattle to review the 2013 assessments of the stocks of aurora and rougheye rockfishes (*Sebastes aurora* and *S. aleutianus*) that are fished off the west coast of the United States. The assessments, and related material, were presented to the Panel; some additional analyses were requested, carried out, and discussed; and the assessment base runs were revised.

I conclude that both assessments are sound and consider them to be the best scientific information available.

I recommend that

- Consideration be given to holding an off-year science workshop on the analysis of survey data, with the aim of fully documenting an agreed approach to this topic.

- Consideration be given to holding an off-year science workshop on the weighting of composition data in stock assessment models.

- Standard software be developed to construct survey and fishery length compositions and to estimate input sample sizes by bootstrapping.

- Priority be given to compiling catch reconstructions for Washington, as have been done for Oregon and California, and that all catch reconstructions should characterize the uncertainty in these catches.

- Consideration be given to making the way Stock Synthesis deals with ‘extra SDs’ for abundance indices more statistically sound.

- Attempts be made to validate the ageing of both species, and that ageing-error measures in future assessments include between-reader error.

1. Background

This report reviews, under contract with the Center for Independent Experts (see Appendix 2), the 2013 assessments of the stocks of aurora and rougheye rockfishes (*Sebastes aurora* and *S. aleutianus*) that are fished off the west coast of the United States. The author was provided with the draft assessment reports and other supporting documents (Appendix 1) and participated both in the meeting which reviewed the assessments, and in the writing of the Panel Report on each assessment.

2. Review Activities

The STAR (stock assessment review) Panel met 8-12 July 2013 at the Northwest Fisheries Science Center (NWFSC) of NOAA/NMFS in Seattle, Washington. Those attending the meeting included four technical reviewers (including the Chair), three advisors, and the stock assessment teams (STATs) (Appendix 3). The assessments, and related material, were presented to the Panel, some additional analyses were requested and carried out, the Panel discussed the results with the assessment team and started to draft its reports (see Annex 3 of Appendix 2 for the agenda).

3. Findings

My findings are grouped according to the seven Terms of References (TORs) for the review, as given in Annex 2 of Appendix 2. For each TOR, I first present findings that are common to both assessments and then continue, where appropriate, with those specific to a particular assessment.

3.1 TOR 1: Prior to the meeting

Become familiar with the draft stock assessment documents, data inputs, and analytical models along with other pertinent information (e.g. previous assessments and STAR panel report when available) prior to review panel meeting.

The documents distributed before the meeting (Appendix 1) allowed the Panel to be well informed about almost all important aspects of the two assessments. I was particularly grateful that the assessment reports included the Stock Synthesis control and data files, because these contain details about the model structure and assumptions that are often over-looked – or sometimes hard to find – in the body of assessment reports.

The one salient weakness in these documents was a lack of information about the design of the trawl surveys and the way that data from these surveys were analysed to provide abundance indices for the assessments. Abundance indices are usually the most important type of input data in stock assessments, so in documenting an assessment it is essential that the way they are calculated is well described and justified. Given the model-based (GLMM) approach used in these assessments, I was surprised to see only limited information about (a) the design of the surveys (including the stratification and any intentional between-stratum variation in sampling density); (b) the relationship between the design and model strata; (c) the model structure, i.e., which predictor variables were used and how (fixed or random?, what interactions?); and (d) the effects estimated in the GLMM.

3.2 TOR 2: Technical merits and deficiencies of the data

Discuss the technical merits and deficiencies of the input data and analytical methods during the open review panel meeting.

3.2.1 Analysis of survey data

I am still unsure as to whether the GLMM approach used (in both assessments) to construct abundance indices for the four trawl surveys was the best approach for these surveys and species. The key issues here are (a) whether it is better to use design- or model-based estimators, and (b) if the latter are used (as they were here), how the model should be structured. I did not feel that enough information was presented for me to reach firm conclusions about these issues. This relative lack of information is perhaps understandable, because I gather that an informal consensus has been reached within the PFMC family on both these issues. However, there is certainly not universal agreement on these issues amongst stock assessment scientists world-wide. I lean towards the usual design-based estimator because its assumptions seem more defensible (e.g., the assumption, in a lognormal GLMM, that the standard deviation of the logarithm of positive catch rates should be the same, in all years and strata is certainly very convenient, but I cannot see why it should be true) and, despite considerable effort, I found that I was unable to build a plausible statistical model for New Zealand trawl survey catch rates (Francis 2006). I don't wish to say that model-based estimators should never be used for survey data, but simply that, when they are used, it is reasonable to require some evidence that they are superior (or at least not inferior) to the conventional design-based estimator. Even those that are already convinced that model-based estimators are superior should expect to see evidence that the model used is well structured.

These matters were discussed at some length by the Panel, but not resolved. Here are brief descriptions of some issues that were of concern to me. Year-to-year changes in the design of early triennial surveys suggest that design-based methods might be preferable to model-based for this series. It was not clear that model strata were always obtained by combining design strata with similar sampling intensities. For rougheye, the choice between random and fixed effects for stratum-year terms was based on median deviance rather than the nature of the variable. Although the triennial survey used multiple vessels neither assessment included vessel in their model for this series. The decision to model vessel effect as random for the NWFSC shelf/slope survey seems odd when there was no randomness in the selection of vessels used each year (two vessels participated in all ten surveys, and another two in seven of the ten). When design- and model-based indices were compared (in Figures 15-18 in the aurora report, and Figure 8 for roughyeye) the former were calculated using the model strata (and thus were not truly design-based). The construction of survey length compositions used the model strata, but ignored vessel effects. For rougheye, a substantial difference in scale between design- and model-based indices (see Figure 8 in the draft assessment report) was assumed to be a result of inadequate software documentation (e.g., a lack of clarity about the units for input data); this is of concern because poor documentation could mean that the model structure used in the analysis differed from that intended by a user.

Another issue discussed by the Panel was the possibility that vessel effects might be aliasing for year effects in models with random vessel effects. This possibility was suggested by a plot of vessel effects for positive rougheye catches in the NWFSC shelf/slope survey in which all (vessel) effects were positive in one year. On further reflection, I am not concerned about the possibility of aliasing as long as the decision to treat vessel effect as random was sound (still an open question for this survey). With truly random vessels (and only a few vessels used each year) it is quite plausible that there would be years in which all vessels had above average catching ability, and the model structure should allow for this.

3.2.2 Construction and relative weighting of length compositions

Questions arose about both the construction and relative weighting of length composition data.

For both fishery and survey data, the construction of a length composition for a given year includes scaling up each individual length sample to represent the whole catch from a landing or station. The question arose as to whether this scaling should use the number or weight of fish in each landing or station. It was surprising to learn that there appeared to be no standard software to make this very routine calculation, so that the choice of scalar (number or weight) was effectively up to the individual analyst. I suspect that though number scaling seems more correct, the choice of scalar will rarely have a strong effect on the outcome of an assessment (this assumption was supported when the aurora survey data, which had been [accidentally?] scaled by weight, were rescaled by number with very little effect on the assessment).

The relative weight assigned to each year's data in a time series of length compositions is determined by what are called 'input *N*s': the multinomial sample sizes assigned to each year before iterative reweighting. The method for setting these input *N*s in the present assessments depended on the species and type of data (fishery or survey) but always derived from some measure of the number of samples (the number of tows, trips or port samples) together, for aurora, with the number of fish. This method has two potential weaknesses. First, for fishery data it is not clear that the sample-size metric is consistent over space and time. Systems for sampling landed catches are often complex, specifically tailored to the way that fish are landed and processed in different regions, and subject to change over time. Thus, what is called a 'trip' in one place and time may not be comparable with the usage in another place or time. During the review we sought evidence of a temporal trend in this metric by requesting plots of the ratio of effective sample size (a model output) to input N against time for each length composition data set (a trend was seen only for sexed data from the aurora trawl fishery), but did not investigate spatial differences. A second weakness is that this approach makes no use of the considerable information about variability that is contained within the data for each year. I believe that a much better way of calculating the input *N*s is to quantify this uncertainty using bootstrap resampling of the data for each year (taking care to resample at all levels – e.g., resampling fish within a landing, and landings within a year) following Crone & Sampson (1998).

The preceding comments about input *N*s apply only to the length compositions. For conditional age at length data I support the approach used in both assessments of setting the input *N* for each length bin equal to the actual sample size.

3.2.3 Minimum sample sizes for aurora length compositions

I think that some of the aurora fishery length composition data should have been omitted from the assessment because sample sizes were much too low to be considered representative. It is difficult to devise objective criteria for a minimum acceptable sample size (and such criteria should consider both the number of trips and fish and their spatial distribution – a sample from just one state is inadequate if much of the catch came from another state and, as in both assessments, there are consistent between-state differences in the fish length). However, it is very clear to me that a sample of just one fish (as occurred nine times in Tables 4 & 5 in the aurora draft assessment report) is far from adequate.

3.3 TOR 3: Model assumptions, estimates, and uncertainty

Evaluate model assumptions, estimates, and major sources of uncertainty.

In both assessments most model assumptions were sound and uncertainty was well dealt with. Stock Synthesis, used for both species, is generally an excellent stock assessment tool, but I am concerned about two aspects of this software (see Sections 3.3.1 and 3.3.4).

3.3.1 Weighting of composition data

In both draft assessments composition data were reweighted using the method (deriving from McAllister & Ianelli 1997, and based on the effective sample sizes output by Stock Synthesis) that has become conventional in this region. This method usually gives too much weight to these data because it ignores correlations between different length bins (Francis 2011). During the review meeting it was demonstrated (using plots like Figure 1) that the composition data sets in both assessments were almost all over-weighted, and these data sets were reweighted in new base case assessments using method TA1.8 of Francis (2011), which allows for correlations. The reweighting, which produced significantly different estimates of depletion, resulted in substantial down-weighting of the length data and lesser down-weighting of the conditional age at length data (for example, down-weighting factors – relative to the original weightings – for the rougheye data ranged from 0.06–0.33 for lengths and 0.63–1.00 for age at length). Paradoxically, the reweighted models for both species showed a *decrease* in uncertainty (narrower confidence limits) for some outputs. This seemed to be because the reweighting effectively reduced the conflict between the length and age-at-length data sets. Changes in estimated recruitment deviates seemed to arise because in the reweighted model these parameters were (appropriately) more influenced by age, than by length data.

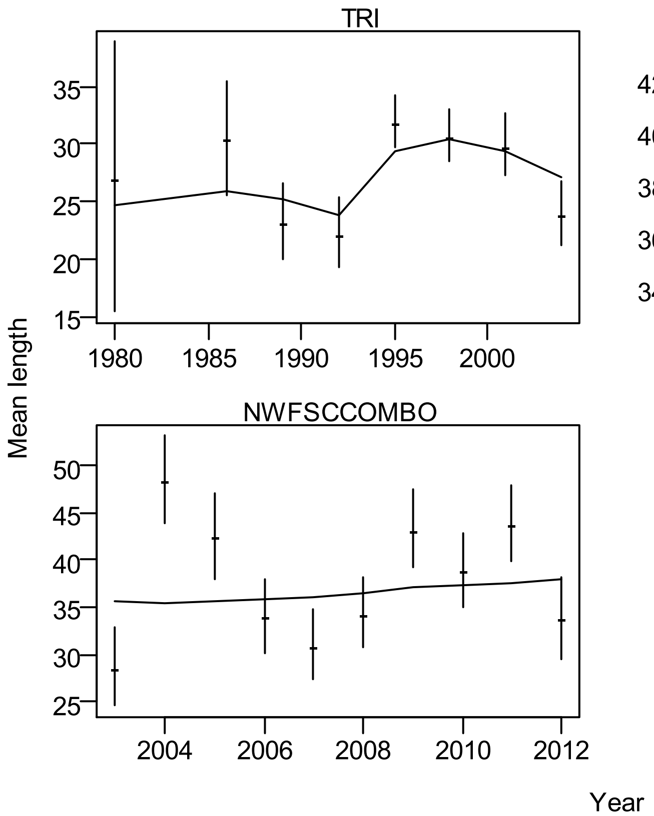

Figure 1: Demonstration that the rougheye survey length composition data sets were over-weighted in the draft base case assessment. Each panel shows the observed ('+', with 95% confidence interval shown as a vertical line) and expected (curved line) mean lengths for one survey. The fact that the expected line often lies outside, or near the edge of, the confidence intervals (whose widths depends on the assumed sample size for each survey and year) indicates that the assumed sample sizes are too large, and thus the data are over-weighted.

For the rougheye assessments, profiles from the reweighted model appeared more satisfactory than those from a run with the original data weighting. For example, the new profile on natural mortality was dominated by the age-at-length data, rather than the length data (which were dominant in the original profile), and the profile on steepness was much flatter than before reweighting (which is consistent with the view that the data contain very little information about this parameter).

3.3.2 Lack of fit to aurora abundance data

I was uncomfortable with the fact that all aurora model runs estimated essentially flat biomass trajectories in their fits to the four abundance indices. None provided fits that were consistent with the trends shown by the indices: a slow increase between 1995 and about 2003, and a slow decline thereafter (Figure 2). Had there been time to explore this lack of fit, I think it would have been useful to determine if is it possible (e.g., by upweighting the abundance indices) to find a model that better fits the trend in Figure 2 and, if so, whether this model estimates a markedly different stock status. Whether such a model should replace the current base case would depend on (a) how hard it was to get the model to fit the abundance indices, and (b) how different the estimated status was.

It may seem unreasonable to ask for better fits to the abundance indices than were found in the base case. I am sure that standard measures of goodness of fit (e.g., the SDNR, or standard deviation of the normalised residuals) would indicate that flat biomass trajectories provide perfectly good fits to the aurora abundance indices, given the c.v.s assigned to them. So, what's the problem? The problem is that these standard measures are based solely on the size of the residuals, but ignore trends in residuals (as is illustrated by figure 5 of Francis 2011). In the aurora assessment they also ignore the fact that the three earliest surveys are consistent in indicating an upward trend. I don't think we should ignore these things.

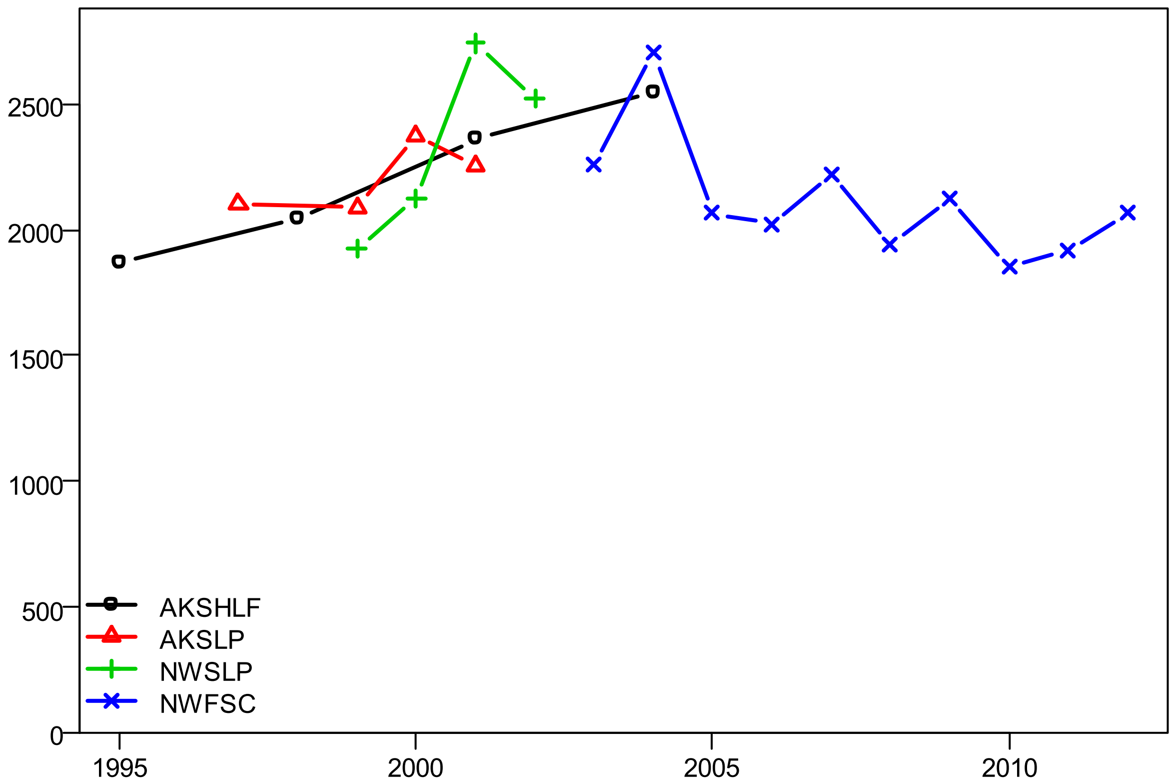

Figure 2: All abundance indices from the aurora rockfish assessment, rescaled to be on similar scales.

3.3.3 Three sources of uncertainty

The main source of uncertainty for both assessments was natural mortality, *M*, and this was well addressed in both draft assessment reports. A closely related topic is ageing uncertainty, because estimation of *M* (either within the model, or as a prior distribution) requires reliable age estimates, and neither species has a validated ageing method. On a related matter, I think that the measures of ageing error used in both assessments were probably under estimates, because neither included between-reader error. There were two other sources of uncertainty that merit mention.

There was considerable uncertainty about the catch histories for both assessments, but particularly for rougheye. It is useful to explore the effects of this uncertainty on the assessment. This was done, to a limited extent, for rougheye during the review and I would like to commend the approach taken. A common approach to this problem, in which the catch uncertainty is spanned by considering alternative low and high (or minimum and maximum) catch histories, is often not very informative, because the effect of these alternative catch histories is mainly to change the scale of the estimated biomass trajectory. What I think can be more informative is to consider a series of alternative hypotheses about aspects of the catch history and see which, if any, of these lead to a material change in the shape of the estimated biomass trajectory. For the rougheye assessment the early hook and line catches were considered uncertain so two alternative catch histories were constructed in which these catches were modified. The fact that these changes had relatively little effect on the assessment output reassured the Panel (and presumably the STAT) that this particular uncertainty was not of great concern. Another obvious area of uncertainty is in the proportion of historic catches in a given market category that is assigned to the assessed species. A consideration of alternative hypotheses about this proportion (which might involve lower or higher constant proportions, or linear temporal trends in the proportion) would be a useful way of determining whether uncertainty about this proportion is material to the assessment. It would be useful if those people reconstructing most likely catch histories for assessed stocks could also document the associated uncertainties in the form of some alternative hypotheses (such as those above).

In the rougheye assessment, one uncertainty which caused much discussion concerned what was called the 'missing teenagers'. This referred to the relative lack of 30-40 cm fish in catches in the triennial and NWFSC shelf/slope surveys which was not reproduced in the model fits. The problem was evident in two plots in the draft assessment report: a plot of fish length against depth for the latter survey (Figure 13A) showed a trend of increasing length with depth and few fish in the 200-300 m depth range where these 'teenagers' were most likely to be found; and aggregate fits to the survey length compositions (Figure 55) showed poor fit in the 30-40 cm range. Additional analyses requested by the Panel failed to resolve this issue.

3.3.4 Estimation of 'extra SDs'

The rougheye assessment used a Stock Synthesis feature which allows the estimation of what are called 'extra SDs' for the abundance indices. I think the way this extra error is implemented in Stock Synthesis is theoretically unsound and misrepresents what is being estimated.

For a given abundance index, the Stock Synthesis approach may be written as $\sigma_{\text{total}, y} = \sigma_{\text{input}, y} + \sigma_{\text{extra}}$, where $\sigma_{\text{total}, y}$ is the s.d. (standard deviation) in log space of the error distribution for the index in year y, $\sigma_{\text{input}, y}$ is an input (user-provided) s.d. for that year, and σ_{extra} is an estimated parameter for that abundance index. This is theoretically unsound, because when statistical errors are added, the size of the total error is obtained by adding variances, not s.d.s. That is, the statistically correct equation is $\sigma_{\text{total},y} = \left(\sigma^2_{\text{input},y} + \sigma^2_{\text{extra}}\right)^{0.5}$.

I suggest it would be better to use the statistical equation. This would remove the misrepresentation mentioned above, as I will illustrate using parameters for the AFSC survey in the rougheye draft assessment report. From Tables 6 and 18 in that report we have σ_{input} = 0.301 and 0.528 for years 1996 and 1997, respectively, and σ_{extra} = 0.045, so the estimated σ_{total} is 0.346 and 0.573. But if we accept the input and total s.d.s as being correct and calculate what σ_{extra} should be, according to the statistical equation, we get values of 0.171 and 0.223, for 1996 and 1997 respectively. In the assessment the additional error is misrepresented as being much smaller than it really is: the estimated s.d. of 0.045 should be about 0.20. This example also makes it clear that actual extra error had a different size in different years.

A change of nomenclature would also be useful in helping users understand the meaning of this feature in Stock Synthesis. For each abundance index there are three separate versions of the abundance in year y: the observed abundance, O_y; the true abundance, T_y; and the model's estimated abundance, E_y (for simplicity I am ignoring the catchability constant). The total error, which is modelled in the likelihood, measures the likely size of the residual $O_y - E_y$. This error is the sum of two parts: the observation error, $O_y - T_y$, and the process error, $T_y - E_y$ (which arises because the model is only an approximation to the real word). Thus, what are called the input and extra s.d.s above might more informatively be called the observation and process error s.d.s.

3.3.5 Two minor issues

There are two other model assumptions I would like to mention, though both are rather minor and are unlikely to have much effect on the assessment outputs.

I think it would be better not to have estimated the initial (1916) age structure in the rougheye assessment. The decision to do so was justified by the desire to include uncertainty about this structure in the model outputs. We didn't accept this argument as justification for estimating steepness, and I don't see why we should for the initial age structure. For aurora, the decision to combine male with female length data below 20 cm did not seem helpful; it simply made the plots of fits to the composition data hard to interpret. I can see that this feature of Stock Synthesis would be useful if the observed sex ratios of the smaller fish were unusually uncertain, but that didn't seem to be the case here.

3.4 TOR 4: Current improvements

Provide constructive suggestions for current improvements if technical deficiencies or major sources of uncertainty are identified.

The main improvement suggested during the review was to improve the weighting of composition data (see Section 3.3.1). Consideration of alternative rougheye catch histories helped to explore uncertainty about catches (see Section 3.3.3) and a discussion of the construction of length compositions improved the scaling of these data in the aurora assessment (see Section 3.2.2).

3.5 TOR 5: Best available science

Determine whether the science reviewed is considered to be the best scientific information available.

I believe both assessments, as modified during the review, represent the best scientific information available. Although I have some reservations about the construction of the abundance indices (see Section 3.2.1), I think it unlikely that a different method of construction would have substantially altered the results of either of these particular assessments.

3.6 TOR 6: Future improvements

When possible, provide specific suggestions for future improvements in any relevant aspects of data collection and treatment, modeling approaches and technical issues, differentiating between the short-term and longer-term time frame.

In considering suggestions for future improvements I have focussed almost exclusively on topics which are generic to stock assessment, since this is where my expertise lies.

3.6.1 Analysis of survey data

Lengthy Panel discussions on this topic suggested to me that although there seemed to be some consensus within the PFMC family about the best approach to the analysis of survey data, neither this consensus nor the available software was well documented, so that the STATs lacked clear guidance on the best approaches and their justification (see issues listed in Section 3.2.1). This seems to me an ideal topic for an off-year science workshop. Because of its importance in many assessments it is an issue that should be addressed in the short term.

3.6.2 Weighting of composition data

I think there is clear evidence that the data-weighting approach most commonly used with Stock Synthesis typically assigns too much weight to composition data (see Section 3.3.1), which compromises model estimates of uncertainty, invalidates statistical inference (e.g., use of AIC [Akaike, 1974] to justify additional model parameters, and construction of profile confidence intervals), and can cause poor fit to abundance indices. There is a need to reach consensus about a better approach, and I suggest that an off-year science workshop might be a sensible way to achieve that consensus. Because of its importance in all assessments this is an issue that should be addressed in the short term

3.6.3 Construction and relative weighting of length compositions

The construction of survey and fishery length compositions is a routine part of many stock assessments so it was surprising to me that no standard software was available for this task. Such software would avoid the inadvertent use of different scaling methods (as occurred in these assessments). Such software should include a bootstrapping facility to allow the user to estimate initial *N*s based on the uncertainty in the actual data, rather than on simple descriptive statistics, such as the number of tows or trips (see Section 3.2.2). Because of its importance in many assessments this is an issue that should be addressed in the short term.

3.6.4 Catch histories

There is a need for catch reconstructions for Washington, as have been compiled for Oregon and California. All catch reconstructions would be enhanced by including, as well best estimates of catches, a set of alternative hypotheses characterizing the uncertainty in these catches (see Section 3.3.3). This work is time consuming and perhaps not as urgent as the preceding issues.

3.6.5 Stock Synthesis modification

The way Stock Synthesis deals with ‘extra SDs’ for abundance indices could be made more statistically sound (see Section 3.3.4). This change would usually have only a slight effect on a stock assessment but it would require only a small modification to the software and documentation.

3.6.6 Ageing

There is an obvious need to validate the ageing of aurora and rougheye in order to reduce uncertainty about natural mortality (the major axis of uncertainty for both species). I note that in other long-lived species (e.g., orange roughy) the break and burn technique used for aurora and rougheye otoliths has been found to be inferior to the use of thin sections. This work is time consuming and not as urgent as the preceding issues. Measures of ageing-error in future assessments would be more realistic if they included between-reader error.

3.7 TOR 7: Panel proceedings

Provide a brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations.

The review process was very well run, and the Panel was well supported and ably chaired. I was impressed by the willingness, and ability, of the assessment teams to respond to panel requests, and was grateful for helpful and constructive comments given by other meeting participants. There were no significant areas of disagreement either within the Panel, or between the Panel and other participants in the review. Major topics of discussion included the analysis of survey data (see Section 3.2.1), construction and weighting of composition data (see Sections 3.2.2 and 3.3.1), and, for rougheye, uncertainties associated with the catch history and the 'missing teenagers' (see Section 3.3.3).

4. Conclusions and recommendations

This was a well run review which I think has strengthened both the assessments and our understanding of them.

I recommend that:

- Consideration be given to holding an off-year science workshop on the analysis of survey data, with the aim of fully documenting an agreed approach to this topic (see Sections 3.2.1, 3.6.1).
- Consideration be given to holding an off-year science workshop on the weighting of composition data in stock assessment models (see Sections 3.3.1, 3.6.1).
- Standard software be developed to construct survey and fishery length compositions and to estimate input sample sizes by bootstrapping (see Sections 3.2.2, 3.6.3).
- Priority be given to compiling catch reconstructions for Washington, as have been done for Oregon and California, and that all catch reconstructions should characterize the uncertainty in these catches (see Sections 3.3.3, 3.6.4).
- Consideration be given to making the way Stock Synthesis deals with 'extra SDs' for abundance indices more statistically sound (see Section 3.3.4).
- Attempts be made to validate the ageing of both species, and that ageing-error measures in future assessments include between-reader error (see Section 3.3.3).

5. References

Akaike, A. 1974. A new look at the statistical model identification. IEEE Trans. Automat. Contr. 19(6): 716–723.

Crone, P.R.; Sampson, D.B. 1998. Evaluation of assumed error structure in stock assessment models that use sample estimates of age composition. In Fishery Stock Assessment Models. Edited by F. Funk, T.J. Quinn, J. Heifetz, J.N. Ianelli, J.E. Powers, J. F. Schweigert, P.J. Sullivan, and C.-I. Zhang. Alaska Sea Grant College Program Report No. AK-SG-98–01. University of Alaska, Fairbanks, Alaska. pp. 355–370.

Francis, R.I.C.C. (2006). Optimum allocation of stations to strata in trawl surveys. *New Zealand Fisheries Assessment Report 2006/23*. 50 p.

Francis, R.I.C.C. (2011). Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences 68: 1124-1138.

McAllister, M.K., and Ianelli, J.N. 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. Can. J. Fish. Aquat. Sci. 54(2): 284–300.

Appendix 1 Materials Provided

The materials provided prior to the review fell into five categories: meeting materials, draft assessment reports, background documents, manuscripts related to rare-catch events, and Stock Synthesis documentation.

Meeting Materials:

1) STAR Panel Meeting Agenda

2) List of Participants Stock Assessment Review Panel

3) Terms of Reference for the Groundfish and Coastal Pelagic Species Stock Assessment and Review Process for 2013-2014. Pacific Fishery Management Council. November, 2012.

4) NOAA's Information Quality Act, Conflict of Interest Policy and OMB's Peer Review Guidelines

Draft Stock Assessment Documents:

Hicks, A. Wetzel, C. and Harms, J. 2013. The status of rougheye rockfish (*Sebastes aleutianus*) and blackspotted rockfish (*S. melanostictus*) as a complex along the U.S. West Coast in 2013 . DRAFT (Pre-STAR) version.

Hamel, O.S., Cope, J.M., and Matson, S. 2013. Stock Assessment of Aurora Rockfish in 2013. DRAFT (Pre-STAR) version.

Background Materials

Hamel, O. Development of prediction intervals and priors for the natural mortality rate using multiple meta-analyses using life-history correlates. NOAA Fisheries, Northwest Fisheries Science Center, Seattle. 4/28/2013.

Karnowski, M, Vladlena Gertseva, and Andi Stephens. 2012. Historical Reconstruction of Oregon's Commercial Fisheries Landings. September, 2012.

NWFSC Observer Program. 2013. Data Products for Stock Assessment Authors. *8Jan. 2013.*

Punt, A.E., Smith, D.C., KrusicGolub, K. and Robertson, S. 2008. Quantifying age-reading error for use in fisheries stock assessments, with application to species in Australia's southern and eastern scalefish and shark fishery. Can. J. Fish. Aquat. Sci. 65: 1991–2005.

Ralston, S., Pearson, D., Field, J., and Key, M. 2009. Documentation of the California Catch Reconstruction Project. April 20, 2009.

Thorson, J. Estimating a Bayesian prior for steepness in Pacific rockfishes (*Sebastes* spp.) off the U.S. West Coast for the 2013 assessment cycle. April 1, 2013.

Thorson, J. T. and Ward, E. Accounting for space-time interactions in index standardization models.

Wallace, J. R. Applying the U.S. West Coast's First Major Trawl Bycatch and Mesh Size Studies to Fishery data using Post-hoc Fishing Strategies and Geographical Area. DRAFT.

Rare Catch Events-Related Manuscripts

Thorson, J.T., Stewart, I.J., and Punt, A.E. 2012. Development and application of an agent-based model to evaluate methods for estimating relative abundance indices for shoaling fish such as Pacific rockfish (Sebastes spp.). ICES Journal of Marine Science, 69(4), 635–647. doi:10.1093/icesjms/fss003.

Thorson, J.T., Stewart, I.J., and Punt, A.E. 2011. Accounting for fish shoals in single- and multi-species survey data using mixture distribution models. CJFAS – Proof.

Thorson, J.T. and Ward, E.J. *In press.* Accounting for space-time interactions in index standardization models.

Stock Synthesis Model-Related Documents

Methot, R. D. 2012. User Manual for Stock Synthesis Model Version 3.24f. Updated October 3, 2012. NOAA Fisheries, Seattle, Washington.

Wetzell, C. 2013. Stock Synthesis Technical Description.

Appendix 2: Statement of Work

This appendix contains the Statement of Work, including three annexes, that formed part of the consulting agreement between Northern Taiga Ventures Inc. and the author.

External Independent Peer Review by the Center for Independent Experts

Stock Assessment Review (STAR) Panel for Rougheye Rockfish and Aurora Rockfish

Scope of Work and CIE Process: The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract providing external expertise through the Center for Independent Experts (CIE) to conduct independent peer reviews of NMFS scientific projects. The Statement of Work (SoW) described herein was established by the NMFS Project Contact and Contracting Officer's Technical Representative (COTR), and reviewed by CIE for compliance with their policy for providing independent expertise that can provide impartial and independent peer review without conflicts of interest. CIE reviewers are selected by the CIE Steering Committee and CIE Coordination Team to conduct the independent peer review of NMFS science in compliance the predetermined Terms of Reference (ToRs) of the peer review. Each CIE reviewer is contracted to deliver an independent peer review report to be approved by the CIE Steering Committee and the report is to be formatted with content requirements as specified in **Annex 1**. This SoW describes the work tasks and deliverables of the CIE reviewer for conducting an independent peer review of the following NMFS project. Further information on the CIE process can be obtained from www.ciereviews.org.

Project Description: New assessments will be conducted for rougheye and aurora rockfishes, which are both considered to be "highly vulnerable species" with vulnerability scores of 2.27 and 2.10, respectively (http://www.pcouncil.org/wp-content/uploads/E2b_GMT_RPT_MARCH_2010_BB.pdf). Assessments for these two stocks will provide the basis for the management of the groundfish fisheries off the West Coast of the U.S. and provide the scientific basis for setting OFLs and ABCs as mandated by the Magnuson-Stevens Act. The technical review will take place during a formal, public, multiple-day meeting of fishery stock assessment experts. Participation of external, independent reviewer is an essential part of the review process. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**.

Requirements for CIE Reviewers: Two CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein. One of the CIE reviewers will participate in all STAR panels held in 2013 to provide a level of consistency between the STAR panels. The CIE reviewers shall be active and engaged participants throughout panel discussions and able to voice concerns, suggestions, and improvements while respectfully interacting with other review panel members, advisors, and stock assessment technical teams. The CIE reviewers shall have excellent communication skills in addition to working knowledge and recent experience in fish population dynamics, with experience in the integrated analysis modeling approach, using age-and size-structured models, use of MCMC to develop confidence intervals, and use of Generalized Linear Models in stock assessment models. Each CIE reviewer's duties shall not exceed a maximum of 14 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall conduct an independent peer review during the panel review meeting scheduled in Seattle, Washington during the dates of 8-12 July 2013.

Statement of Tasks: Each CIE reviewers shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

Prior to the Peer Review: Upon completion of the CIE reviewer selection by the CIE Steering Committee, the CIE shall provide the CIE reviewer information (full name, title, affiliation, country, address, email) to the COTR, who forwards this information to the NMFS Project Contact no later the date specified in the Schedule of Milestones and Deliverables. The CIE is responsible for providing the SoW and ToRs to the CIE reviewers. The NMFS Project Contact is responsible for providing the CIE reviewers with the background documents, reports, foreign national security clearance, and other information concerning pertinent meeting arrangements. The NMFS Project Contact is also responsible for providing the Chair a copy of the SoW in advance of the panel review meeting. Any changes to the SoW or ToRs must be made through the COTR prior to the commencement of the peer review.

Foreign National Security Clearance: When CIE reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for CIE reviewers who are non-US citizens. For this reason, the CIE reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/sponsor.html).

Pre-review Background Documents: Two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE Lead Coordinator on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.
Documents to be provided to the CIE reviewers prior to the STAR Panel meeting include:

- The current draft stock assessment reports;
- The Pacific Fishery Management Council's Scientific and Statistical Committee's Terms of Reference for Stock Assessments and STAR Panel Reviews;
- Stock Synthesis (SS) Documentation
- Additional supporting documents as available.
- An electronic copy of the data, the parameters, and the model used for the assessments (if requested by reviewer).

Panel Review Meeting: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs can not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COTR and CIE Lead Coordinator.** Each CIE reviewer shall actively participate in a professional and respectful manner as a member of the meeting review panel, and their peer review tasks shall be focused on the ToRs as specified herein. The NMFS Project Contact is responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements). The NMFS Project Contact is responsible for ensuring that the Chair understands the contractual role of the CIE reviewers as specified herein. The CIE Lead Coordinator can contact the Project Contact to confirm any peer review arrangements, including the meeting facility arrangements.

Contract Deliverables – Independent CIE Peer Review Reports: Each CIE reviewer shall complete an independent peer review report in accordance with the SoW. Each CIE reviewer shall complete the independent peer review according to required format and content as described in Annex 1. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in Annex 2.

Other Tasks – Contribution to Summary Report: Each CIE reviewer may assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review. Each CIE reviewer is not required to reach a consensus, and should provide a brief summary of the reviewer's views on the summary of findings and conclusions reached by the review panel in accordance with the ToRs.

Specific Tasks for CIE Reviewers: The following chronological list of tasks shall be completed by each CIE reviewer in a timely manner as specified in the **Schedule of Milestones and Deliverables**.

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Participate during the panel review meeting in Seattle, Washington during the dates of 8-12 July, 2013 as specified herein, and conduct an independent peer review in accordance with the ToRs (**Annex 2**).
3) No later than 26 July 2013, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Mr.

Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and to Dr. David Die, CIE Regional Coordinator, via email to ddie@rsmas.miami.edu. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.

Schedule of Milestones and Deliverables: CIE shall complete the tasks and deliverables described in this SoW in accordance with the following schedule.

June 3, 2013	CIE sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact
June 24, 2013	NMFS Project Contact sends the CIE Reviewers the pre-review documents
July 8-12, 2013	Each reviewer participates and conducts an independent peer review during the panel review meeting
July 26, 2013	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
August 2, 2013	CIE submits CIE independent peer review reports to the COR
August 9, 2013	The COR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

Modifications to the Statement of Work: Requests to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent substitutions. The Contracting Officer will notify the COTR within 10 working days after receipt of all required information of the decision on substitutions. The COTR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role and ability of the CIE reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and ToRs shall not be changed once the peer review has begun.

Acceptance of Deliverables: Upon review and acceptance of the CIE independent peer review reports by the CIE Lead Coordinator, Regional Coordinator, and Steering Committee, these reports shall be sent to the COTR for final approval as contract deliverables based on compliance with the SoW and ToRs. As specified in the Schedule of Milestones and Deliverables, the CIE shall send via e-mail the contract deliverables (CIE independent peer review reports) to the COTR (William Michaels, via William.Michaels@noaa.gov).

Applicable Performance Standards: The contract is successfully completed when the COTR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards:
(1) each CIE report shall completed with the format and content in accordance with **Annex 1**,
(2) each CIE report shall address each ToR as specified in **Annex 2**,
(3) the CIE reports shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

Distribution of Approved Deliverables: Upon acceptance by the COTR, the CIE Lead Coordinator shall send via e-mail the final CIE reports in *.PDF format to the COTR. The COTR will distribute the CIE reports to the NMFS Project Contact and Center Director.

Support Personnel:

William Michaels, Program Manager, COTR

NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov Phone: 301-713-2363 ext 136

Manoj Shivlani, CIE Lead Coordinator
Northern Taiga Ventures, Inc.
10600 SW 131st Court, Miami, FL 33186
shivlanim@bellsouth.net Phone: 305-383-4229

Roger W. Peretti, Executive Vice President
Northern Taiga Ventures, Inc. (NTVI)
22375 Broderick Drive, Suite 215, Sterling, VA 20166
RPerretti@ntvifederal.com Phone: 571-223-7717

Key Personnel:

Stacey Miller, NMFS Project Contact
National Marine Fisheries Service,
2032 SE OSU Drive,
Newport OR 97365
Stacey.Miller@noaa.gov
Phone: 541-867-0562

Michelle McClure
National Marine Fisheries Service,
2725 Montlake Blvd. E,
Seattle WA 98112
Michelle.McClure@noaa.gov

Jim Hastie
National Marine Fisheries Service,
2725 Montlake Blvd. E,
Seattle WA 98112
Jim.Hastie@noaa.gov
Phone: 541-867-3412

Appendix 2, Annex 1: Format and Contents of CIE Independent Peer Review Report

1. The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

   a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including providing a brief summary of findings, of the science, conclusions, and recommendations.

   b. Reviewers should discuss their independent views on each ToR even if these were consistent with those of other panelists, and especially where there were divergent views.

   c. Reviewers should elaborate on any points raised in the Summary Report that they feel might require further clarification.

   d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

   e. The CIE independent report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The CIE independent report shall be an independent peer review of each ToRs, and shall not simply repeat the contents of the summary report.

3. The reviewer report shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of the CIE Statement of Work
   Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

Appendix 2, Annex 2: Terms of Reference for the Peer Review

Stock Assessment Review (STAR) Panel for Aurora and Rougheye Rockfishes

1. Become familiar with the draft stock assessment documents, data inputs, and analytical models along with other pertinent information (e.g. previous assessments and STAR panel report when available) prior to review panel meeting.
2. Discuss the technical merits and deficiencies of the input data and analytical methods during the open review panel meeting.
3. Evaluate model assumptions, estimates, and major sources of uncertainty.
4. Provide constructive suggestions for current improvements if technical deficiencies or major sources of uncertainty are identified.
5. Determine whether the science reviewed is considered to be the best scientific information available.
6. When possible, provide specific suggestions for future improvements in any relevant aspects of data collection and treatment, modeling approaches and technical issues, differentiating between the short-term and longer-term time frame.
7. Provide a brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations.

Appendix 2, Annex 3: Tentative Agenda

Final Agenda to be provided two weeks prior to the meeting with draft assessments and background materials.

Stock Assessment Review (STAR) Panel for
Aurora and Rougheye Rockfishes

NMFS Northwest Fisheries Science Center
Auditorium
2725 Montlake Blvd, NE
Seattle, WA 98112

July 8-12, 2013

Monday, July 8, 2013

8:30 a.m.	Welcome and Introductions
9:15 a.m.	Review the Draft Agenda and Discuss Meeting Format (D. Sampson, Chair) - Review Terms of Reference (TOR) for assessments and STAR panel - Assign reporting duties - Discuss and agree to format for the final assessment document - Agree on time and method for accepting public comments
9:30 a.m.	Presentation of Aurora Rockfish Assessment (A. Hicks) - Overview of data and modeling
12:30 p.m.	Lunch (On Your Own)
1:30 p.m.	Q&A session with Aurora rockfish STAT STAR Panel discussion - Panel develops written request for additional model runs / analyses
3:30 p.m.	Presentation of Rougheye Rockfish Assessment (O. Hamel) (if time allows) - Overview of data and modeling
5:30 p.m.	Adjourn for Day.

Tuesday, July 9, 2013

8:30 a.m.	Continue Presentation of Rougheye Rockfish Assessment (O. Hamel) - Overview of data and modeling
12:00 p.m.	Lunch (On Your Own)
1:30 p.m.	Q&A Session with Rougheye Rockfish STAT Panel Discussion - Panel develops written request for additional model runs / analyses
4:30 p.m.	Check in with Aurora rockfish -STAT
5:30 p.m.	Adjourn for Day.

Wednesday, July 10, 2013

8:30 a.m.	Presentation of First Set of Model Runs for Aurora Rockfish (A. Hicks) - Q&A session with the Aurora Rockfish & Panel discussion - Panel develops written request for second round of model runs / analyses for aurora rockfish -STAT

12:00 p.m.	Lunch (On Your Own)
1:30 p.m.	Presentation of First Set of Model Runs for Rougheye Rockfish (O. Hamel) - Q&A session with rougheye rockfish -STAT & panel discussion - Panel develops written request for second round of model runs / analyses for rougheye rockfish -STAT.
5:30 p.m.	Adjourn for day.

Thursday, July 11, 2013

8:30 a.m.	Presentation of Second Set of Model Runs for Aurora Rockfish (A. Hicks) - Q&A session with the aurora rockfish-STAT & panel discussion - Agreement of preferred model and model runs for decision table - Panel continues drafting STAR report.
12:00 p.m.	Lunch (On Your Own)
1:00 p.m.	Presentation of Second Set of Model Runs for Rougheye Rockfish (O. Hamel) - Q&A session with the rougheye rockfish -STAT & panel discussion - Agreement of preferred model and model runs for decision table - Panel continues drafting STAR report.
4:00 p.m.	Continue Panel Discussion or Drafting STAR Panel Report
5:30 p.m.	Adjourn for day.

Friday, July 12, 2013

8:30 a.m.	Consideration of Remaining Issues - Review decision tables for assessments
10:00 a.m.	Panel Report Drafting Session
12:00 p.m.	Lunch (on your own)
2:00 p.m.	Review First Draft of STAR Panel Report
4:00 p.m.	Panel Agrees to Process for Completing Final STAR Report by Council's September Meeting Briefing Book Deadline
5:30 p.m.	Review Panel Adjourn.

Appendix 3: Panel Membership

The review panel comprised four technical reviewers (including the Chair), three advisors, and the stock assessment teams.

Technical Reviewers
David Sampson Scientific and Statistical Committee (SSC), Panel Chair
Yan Jiao, Center for Independent Experts (CIE)
Chris Francis, Center for Independent Experts (CIE)
John Field, Southwest Fisheries Science Center (SWFSC)

Panel Advisors
John DeVore, Pacific Fishery Management Council (PFMC), Staff Officer
Colby Brady, PFMC Groundfish Management Team (GMT)
Gerry Richter, PFMC Groundfish Advisory Subpanel (GAP)

Stock Assessment Teams (STATs)
Rougheye Rockfish STAT
Allan Hicks, Northwest Fisheries Science Center (NWFSC)
Chantell Wetzel, Northwest Fisheries Science Center (NWFSC)
John Harms, Northwest Fisheries Science Center (NWFSC)

Aurora Rockfish STAT
Owen Hamel, Northwest Fisheries Science Center (NWFSC)
Jason Cope, Northwest Fisheries Science Center (NWFSC)
Sean Matson, Northwest Regional Office (NWRO)